Data Compression - Systematisation

---- Overview ----

by T.Strutz

My investigations led to the result that all techniques and methods of data compression can be subdivided into three categories (see figure 1).
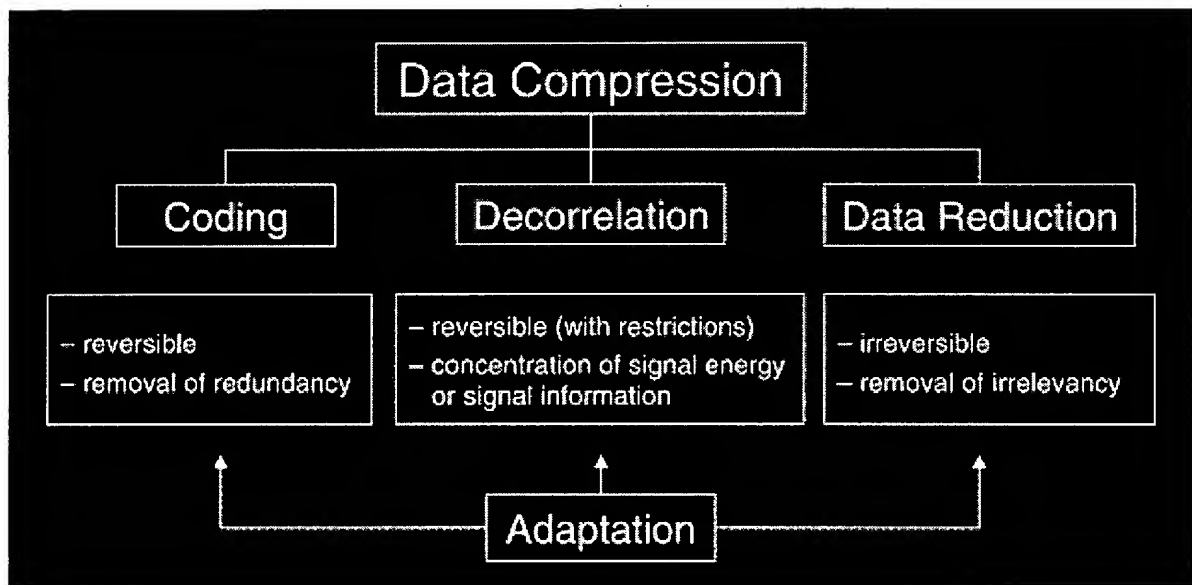

Figure 1:

Coding comprises all techniques aiming to a reduction of redundancy in the signal. The amount of redundancy inside a signal depends on the distribution of the signal values and the statistical bindings between them. Thus, uniform distributed white noise, for instance, contains no redundancy. Coding techniques are reversible, which means the operations can be inverted without loss of digital information (identity between original and reconstructed signal).

Data reduction comprises all methods, which reduce the irrelevant parts of the signal content. This results into changed signal information.

These both categories are supplemented by a third, which I have termed **decorrelation**. All techniques that concentrate the signal information (or signal energy) to few signal values, belong to this class. Typically, the distribution of the signal values becomes sharper by decorrelation. Most of the decorrelation algorithms work with floating point numbers. In this case the operations are not reversible due to the limited resolution of digital numbers in computers. Some methods can operate with integer numbers and guarantee reversibility.

Generally, all processing steps are based on assumptions at the signal in terms of a signal model and sets of parameters. In case of changing parameters or model alterations the compression result becomes worse. This can be avoided by **adaptation**. The analysis of existing compression systems or the

h   e ech   c e   c e   c e   h

development of new ones should always distinguish between the basic technique and its adaptation.

back to Start

---

h   e ech   c   e   c   e   c   e   h